

Accurate genotyping of hepatitis C virus through nucleotide sequencing and identification of new HCV subtypes in China population

Y.-Q. Tong^{1,2}, B. Liu³, H. Liu⁴, H.-Y. Zheng², J. Gu¹, H. Liu², E. J. Song⁵, C. Song⁵ and Y. Li^{1,2}

1) Department of Clinical Laboratory, 2) Clinical Molecular Diagnostic Centre, Renmin Hospital of Wuhan University, 3) Department of Pathology, Affiliated Tianyou Hospital of Wuhan University of Science and Technology, Wuhan, 4) Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai, China and 5) Pennsylvania State University College of Medicine and Hershey Medical Center, Hershey, PA, USA

Abstract

Nucleotide sequencing of the phylogenetically informative region of NS5B remains the gold standard for hepatitis C virus (HCV) genotyping. Here we developed a new methodology for sequencing new NS5B regions to increase the accuracy and sensitivity of HCV genotyping and subtyping. The eight new primers were identified by scanning the full-length NS5B regions from 1127 HCV genomic sequences found in HCV databases. The ability of each pair of primers to amplify HCV subtypes was scored, and the new primers were able to amplify the NS5B region better than the previously used primers, therefore more accurately subtyping HCV strains. Sequencing the DNA amplified by the new primer pairs can specifically and correctly detect the five standard HCV subtypes (1a, 2a, 3b, 6a and 1b). We further examined patient samples and found that the new primers were able to identify HCV subtypes in clinical samples with high sensitivity. This method was able to detect all subtypes of HCV in 567 clinical samples. Importantly, three novel HCV subtypes (1b-2a, 1b-2k and 6d-6k) were identified in the samples, which have not been previously reported in China. In conclusion, sequencing the NS5B region amplified by the new NS5B primers is a more reliable method of HCV genotyping and a more sensitive diagnostic tool than sequencing using the previously described primers, and could identify new HCV subtypes. Our research is useful for clinical diagnosis, guidance of clinical treatment, management of clinical patients, and studies on the epidemiology of HCV. Clinical Microbiology and Infection © 2015 European Society of Clinical Microbiology and Infectious Diseases. Published by Elsevier Ltd. All rights reserved.

Keywords: Genotype, hepatitis C virus, ns5b, sequencing, subtype

Original Submission: 7 January 2015; **Revised Submission:** 10 May 2015; **Accepted:** 26 May 2015

Editor: G. Antonelli

Article published online: 6 June 2015

Corresponding authors: Y. Li, Department of Clinical Laboratory, Renmin Hospital of Wuhan University, 99 Ziyang Road of Wuchang District, Wuhan 430060, China

C. Song, Pennsylvania State University College of Medicine, Penn State Hershey Children's Hospital, PO Box 850, 500 University Drive, Hershey, PA 17033, USA

E-mails: csong@hmc.psu.edu (C. Song), climdc@163.com (Y. Li)

Y.-Q. Tong, B. Liu and H. Liu are contributed equally to this work.

Introduction

Hepatitis C virus (HCV) is a blood-borne pathogen that is spread through blood and other body fluids [1,2]. It is among

the most common viruses infecting the liver; every year, 3–4 million people are infected by HCV [3]. HCV infection is only cleared in approximately 25% of the cases, leaving about 150 million people chronically infected [4]. Each year more than 350 000 people die from HCV-related liver diseases, such as chronic liver disease, cirrhosis and hepatocellular carcinoma (HCC) [5].

Hepatitis C virus is a positive-sense RNA virus with a genome of about 9400 bases, which encodes a single poly-protein that is cleaved into four structural proteins (Core, Envelope 1 and 2, and p7) and six non-structural proteins named NS2–NS5B [6]. HCV is divided into six distinct genotypes throughout the world with multiple subtypes in each genotype class [7–9]. Variations in the HCV genome are used to classify the virus into genotypes [10]. Accurate classification of genotypes and subtypes of HCV is important for correct

stratification of groups in clinical trials and accurate analysis of data related to efficacy and resistance of new HCV drugs. It is also essential for the implementation of therapeutic procedures, the production of effective vaccines, the improvement of diagnostic tests [11–13], and a better understanding of the epidemiological and physical features of the virus [14].

Importantly, the tendency for development of HCC is increased among people with chronic HCV infection [15,16]. Also, the rate of HCV infection is unexpectedly high in the Asian population, and HCV infection at an older age may give a greater risk of HCC [17]. Therefore, the development of a more sensitive and accurate methodology for HCV genotyping would have a significant impact on further understanding the correlation of HCV genotypes with evolution of severe liver diseases, particularly HCC.

Many genotyping methods targeting different regions of the HCV genome have been developed, such as restriction fragment length polymorphism, line probe assay, TaqMan PCR, liquid microarray, sequencing, solid-phase electrochemical array [18–21]. Sequencing of a genome region divergent enough to discriminate genotypes and subtypes, such as HCV NS5B, core and envelope regions, is considered to be the most accurate method [22,23]. This method of genotyping the HCV 5'-untranslated region gives sufficient information for clinical purposes, where the determination of the subtype is not essential. NS5B genotyping methods are more reliable for subtype determination [24,25]. Direct sequencing of a portion of the NS5B gene followed by phylogenetic analysis is the standard method for identifying HCV genotypes and subtypes [26]. To date, sequencing the NS5B gene to analyse HCV genotype and subtype is considered the most reliable method; however, most reports have used a partial short NS5B fragment to genotype HCV [27,28], and such a classification has certain drawbacks such as low accuracy and a limit to the number of subtypes it can identify. In this study, all nucleic acid sequences from the HCV database were used to predict potential primers located in the NS5B region by PRIMER 5.0, and new fragments of the NS5B gene were amplified by the resulting primers and sequenced for HCV genotyping. The primers can correctly amplify the known nucleic acid sequences of standard HCV strains. HCV was also accurately detected by these primers in patient samples. Results demonstrate that this method can accurately detect the subtypes of all HCV patient samples with high sensitivity.

Materials and methods

Ethics statement

This study was approved by the Institutional Review Board of Renmin Hospital, Wuhan University School of Medicine. A

written informed consent was obtained from each participant in accordance with the Renmin Hospital of Wuhan University Ethics Committee.

HCV sequence databases and PCR primer design

Hepatitis C virus sequence databases (<http://hcv.lanl.gov>) were downloaded from the genome databases of the University of Alabama at Birmingham (<ftp://ftp.genome.uab.edu/vbrc>). All 1127 nucleotide sequences of HCV NS5B from the database were clustered with MEGA 5.0 software. The clustered sequences of NS5B region were used to search all possible forward primers and reverse primers with PRIMER 5.0. Next, possible matching pairs of primers were searched in all resulting primers. If the pair of primers was able to hybridize to the sequence and met the requirements of common primer design, we determined if the pair of primers could amplify the fragment of the NS5B region. With these strict criteria, the best primers were identified. Then, we examined whether each pair of primers matches the corresponding site in each NS5B region from all 1127 HCV sequences by computing and calculated the ratio of each pair of primers for each HCV subtype in the 1127 HCV sequences and the sensitivity for amplification. As control, one pair of the reported primers was also used for this analysis. Then, we did phylogenetic analysis of all the amplified NS5B sequences by the pairs of primers and compared that to the reported pair of primers.

Patients and plasma preparations

A total of 567 samples of 3–5 mL EDTA-anticoagulated peripheral blood were obtained from individuals with HCV infections confirmed by quantitative real-time RT-PCR before treatment in the Department of Infectious Diseases, Renmin Hospital of Wuhan University. The HCV RNA titres of all patients in this cohort were $>10^3$ IU/mL in the peripheral blood with the conventional PCR amplification method. For HCV RNA titres of the patients $<10^3$ IU/mL, a more sensitive method is required. We developed a nested PCR method using universal primers, which can detect HCV RNA titres in 80 IU/mL (data not shown). The HCV patients included patients with acute hepatitis C, chronic hepatitis C, hepatitis C cirrhosis and HCV acute exacerbation with no anti-viral treatment given while they received HCV genotyping testing. The patients' ages ranged from 21 to 65 years (median 41.3 years). All samples were centrifuged for 5 min at 3000 g, and the supernatants were collected and stored at -70° . Five HCV standard strains including HCV 1a, 2a, 3b, 6a and 1b were purchased from the Institute of Virology, Wuhan University.

HCV RNA extraction

Viral RNA was extracted from plasma using the Viral RNA Mini Spin Kit following the manufacturer's instructions (Qiagen,

Hilden, Germany). Briefly, the plasma volume was adjusted based on viral load, i.e. 140 µL for HCV RNA ≥ 5000 IU/mL and 280 µL for HCV RNA < 5000 IU/mL, and mixed with protease K (40–120 µL) and AVL buffer (560–1120 µL) supplemented with 5.6–11.2 µg of carrier RNA (tRNA). The mixture was incubated for 10 min at room temperature, and absolute ethanol (560–1120 µL) was added to each sample and mixed by pulse-vortexing for 15 s. The mixture was then loaded onto a mini column and centrifuged at 6000 g. Subsequently, the mini column was washed with AW1 buffer (500 µL) and AW2 buffer (500 µL), respectively with centrifugation at 6000 g for 1 min. A second wash with AW2 buffer (500 µL) was carried out with centrifugation at 20 000 g for 3 min, and the column was dried with full speed centrifugation for 1 min. The RNA was eluted by loading 50 µL of Buffer AVE onto each column, incubating for 1 min at room temperature and centrifuging at 6000 g for 1 min. The elution was repeated once to increase the yield. Finally, a total of 50 µL viral RNA was isolated from 140–280 µL of plasma.

cDNA synthesis

Complementary DNA was synthesized with the Thermo Scientific RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA) from HCV RNA following the manufacturer's instructions. Briefly, HCV RNA was diluted 1:2 or 1:4 based on the viral load, i.e. 1:2 for HCV RNA < 5000 IU/mL and 1:4 for HCV RNA ≥ 5000 IU/mL, and mixed with a total RT reaction volume of 20 µL composed of 1 µL of a random hexamer primer, 200 units of RevertAid M-MuLV reverse transcriptase, 20 units of RiboLock RNase inhibitor, 4 µL of 5× reaction buffer, and 4 µL of 10 mM dNTP mix. The RT reaction was carried out in a thermocycler with the following conditions: 25°C for 5 min, 65°C for 5 min, 42°C for 60 min, and 70°C for 5 min.

Amplification of the HCV NS5B gene fragments by PCR

The HCV cDNA template was used for the PCR amplification, and the PCR was performed as follows: 94°C for 3 min; 94°C for 30 s, 56°C for 40 s, 72°C for 60 s, 35 cycles; 72°C for 10 min. The primers were selected in the same range of annealing temperature as far as possible. The average T_m for all selected primers was 56–60°C. Multiple pre-experiments with different annealing temperature from 54°C to 62°C were performed, and the best annealing temperature was finally defined as 56°C, which could efficiently amplify all the fragment with only slightly reduction of efficiency rate for F1, R4, F3 and R4. The resulting PCR products were electrophoresed on 1.2% agarose gel for gel purification and documented by imaging. The primers for amplifying NS5B gene fragments are shown in Table 1. The conserved regions of 3' non-coding region of HCV

TABLE 1. The predicted primers for amplifying NS5B gene fragments

Primer name	Primer sequence (5' to 3')	Location (nucleotide)
F1	TTAACCACATCMRCTCCGTGTG	347–368
F2	ACACCCGCTGYTTYGACTC	659–677
F3	GCCTTCACGGARGCYATG	1012–1029
F4	CATGCCTCAGRAARCTTGG	1460–1478
R1	GTCAAAGCAGCGGGTRTCRTA	675–655
R2	GTACCTGGTCATAGCYTCCGTRAA	1038–1015
R3	GGGACCCCAAGYTTYCTGAG	1484–1465
R4	GCCGGCATTGGAGTGAGYTT	1622–1603

Degenerate base symbol: M, A and C; R, A and G; Y, C and T.

were used as control, which was amplified with forward primer 5'-GCGGAACCGGTGAGTACA-3' and reverse primer 5'-CCTATCAGGCAGTACCACAAGG-3'.

DNA purification and sequencing

The gel-purified PCR products were used for sequencing with the dideoxynucleotide chain termination method with the ABI PRISM® BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA, USA). Sequence analysis was performed with the ABITM3130 Genetic Analyzer (Applied Biosystems).

Phylogenetic analysis

All nucleotide sequences from HCV strains were aligned using CLUSTALW 2.0 software with a reference panel of reported sequences available in the HCV sequence database (<http://hcv.lanl.gov/content/hcv-db/index>) provided by the Los Alamos National Laboratory. Pairwise evolutionary distance matrices for the NS5B nucleotide sequences were computed using the p-distance algorithm of the MEGA software package (version 5.2, 2012; Pennsylvania State University, University Park, PA, USA). Sequence distance matrices were analysed with conventional statistical software (SYSTAT, v13.0; Systat Software, Inc., Point Richmond, CA, USA) using files derived from MEGA distance tables. For NS5B sequences, a p-distance of 0.12 was found to delineate strains of the same subtype and strains of a different subtype. Phylogenetic analysis was performed with the MEGA software package using the p-distance algorithm for distance determination and the neighbour-joining method for tree drawing. The reliability of phylogenetic classification was evaluated by a 1000-replication bootstrap test using the same software package. Genotyping was performed using the nomenclature defined by Simmonds *et al.* [29].

Statistical analysis

The distribution of HCV genotypes was compared using chi-square test and Fisher test. Differences were considered to be statistically significant when $p < 0.05$.

Results

Database analysis for HCV nucleic acid sequence

The HCV sequence databases (<http://hcv.lanl.gov>) were searched for different subtypes of HCV genome. A total of 1127 HCV genome nucleic acid sequences were identified. As reference, we first analysed the distribution of the HCV subtypes in the database, which is shown in Fig. 1a and the number of each subtype is shown in Fig. 1b. From these data, we found

that the main genotypes among these sequences were HCV 1a, 1b, 2b and 4a (Fig. 1a), which respectively accounted for 38.86% (438/1127), 29.72% (335/1127), 2.22% (25/1127) and 1.69% (19/1127) (Fig. 1b). Also there were variants, such as HCV VAR1a, VAR1b and VAR2a (Fig. 1a), which respectively accounted for 2.57% (29/1127), 7.72% (87/1127) and 3.99% (45/1127) (Fig. 1b). Then, we used all nucleic acid sequences in the NS5B region (full-length sequence of NS5B) to perform the phylogenetic analysis for subtyping the 1127 HCV sequences. The subtypes of the 1127 HCV sequences analysed with full-

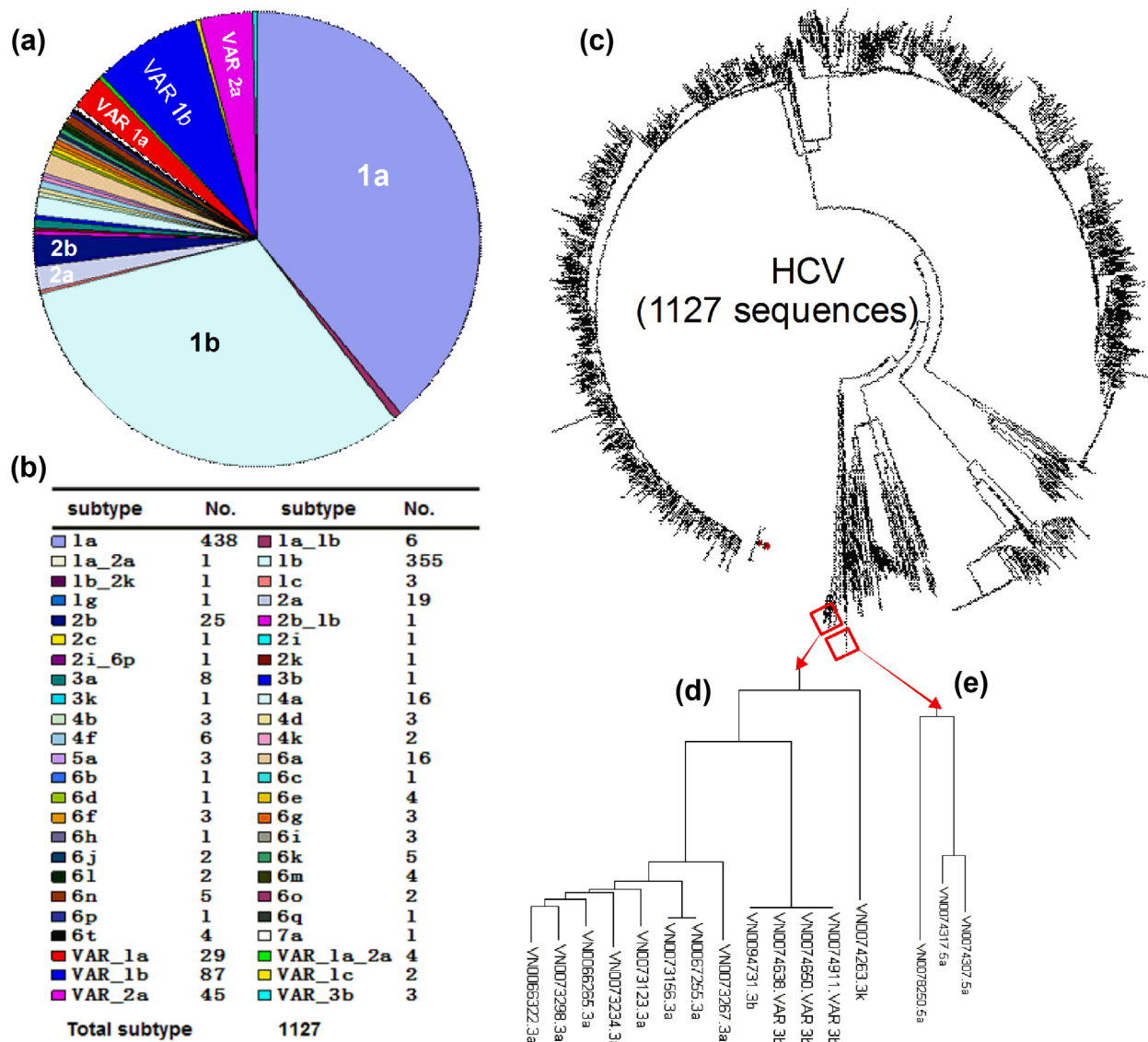


FIG. 1. Analysis of the hepatitis C virus (HCV) type and subtype from Viral Bioinformatics Resource Centre (a, b) the distribution frequency of HCV subtypes in 1127 HCV sequences as shown on a pie chart (a) and numbers of each subtypes as shown (b) in all the HCV sequences. (c–e) Phylogenetic analysis of the NS5B genomic region of HCV G1 strains. Genetic distance was estimated by the Kimura two-parameter model, and a phylogenetic tree was constructed using the neighbour joining method. The phylogenetic analyses of the HCV type 3 (d) and HCV type 5 (e) are further enlarged from (c) to indicate full-length NS5B could clearly distinguish the subtypes of the HCV.

length sequence of NS5B region were the same as those of the reference subtypes for each of the 1127 HCV sequences, and the phylogenetic tree generated with a full-length sequence of the NS5B region was the same as that of the referenced one (Fig. 1c). The two regions of the phylogenetic tree enlarged from Fig. 1c, clearly showed the subtypes of HCV 5 and HCV 3, which is also same as that in the reference database (Fig. 1d, e). These data indicate that the full-length sequence of the NS5B region can correctly classify the subtypes of the HCV.

Furthermore, we analysed the genotypes using the HCV sequences available not only in the above database (<http://hcv.lanl.gov>) but also the NCBI database (<http://www.ncbi.nlm.nih.gov/projects/genotyping>), which contains 32 HCV subtypes including Subtype 1a (AF009606, AF011751, AF271632), Subtype 1b (D90208, AB049088, AF356827, AF139594), Subtype 1c (D14853, AY051292), Subtype 2a (D00944, AF169004, AB047639, AB047641), Subtype 2b (D10988, AF238486, AB030907, AY232730), Subtype 2c (D50409), Subtype 2i (DQ155561), Subtype 2k (AB031663), Subtype 3a (D17763, D28917, AF046866), Subtype 3b (D49374), Subtype 3k (D63821), Subtype 4a (Y11604), Subtype 5a (Y13184, AF064490), Subtype 6a (Y12083), Subtype 6b (D84262), Subtype 6d (D84263), Subtype 6g (D63822), Subtype 6h (D84265), and Subtype 6k (D84264).

Phylogenetic analysis was performed to evaluate the genetic relatedness of the 32 sequences in the NCBI HCV database with the reported nucleic acid sequence from 8370 to 8708 or 8323 to 8638 in the NS5B region. The data showed that analysis with the standard NS5B region mistakenly clustered subtype 5 into subtype 1, and clustered subtype 4 and subtype 5 together (Fig. 2a). We also carried out the phylogenetic analysis with the full-length sequence of the NS5B region and found subtypes 4 and 1 or subtypes 5 and 3 were clustered together but clearly distinct (Fig. 2b), which is consistent with the data analysed with the full-length sequence of NS5B region in 1127 HCV sequences as shown in Fig. 2c. Taken together, these analyses indicate that subtyping with the full-length sequence of NS5B could accurately distinguish the subtypes in both databases but had less accuracy if using the reported partial NS5B region for HCV genotyping. These analyses also suggested that sequencing the full-length NS5B region was more accurate for HCV subtyping.

Primers designed to cover the full-length of NS5B region in all HCV sequences in the database

Each base in the NS5B region of the 1127 HCV sequences was scanned with PRIMER 5.0 software to check suitable primers. The base in the primer population was endowed with increasing weight, which was transformed into image grey-scale. A region with deep colour represents the high PCR amplification ability

across the whole HCV population (Fig. 3A,a,b). A threshold was set for clear image display and pixels less than the threshold (500 dpi) was set as white colour. The oligos sequence must be the same as that of the NS5B sequence or has up to two similar oligonucleotides but these similar nucleotides must not be at the 3' end. All of the 1127 NS5B sequences were initially scanned to calculate the likelihood of each primer pair amplifying the corresponding NS5B genotype, and were subsequently screened for appropriate primer pairs with high efficiency for common genotypes and high sensitivity for rare genotypes in the detection (shown as red boxes in Fig. 3a,b). Based on these criteria, eight primers were identified (Fig. 3c and Table 1), which are able to amplify the region from site 347 to 1603 of NS5B in almost all 1127 HCV subtypes. We also found that all the amplification products of the eight primers are located in the red zones (high intensity, cover more HCV sequences) on Fig. 3a,b. These analyses suggest that the eight primers have high efficiency and sensitivity for amplification.

Next, all possible pairs of the eight primers (ten pairs) were used to examine their sensitivity in amplifying the fragments of the NS5B gene region. The amplification sensitivity of each primer pair in detecting HCV genotypes and subtypes was evaluated by blasting the HCV database. Overall, the pairs of the eight primers could successfully amplify 1117 fragments out of 1127 NS5B, except for seven HCV 1b, one HCV 2b, one HCV 3a and HCV 7 with 88.73–96.45% sensitivity for amplification (Table 2), which is much higher than that with the reported primers, which could not detect two of seven genotypes and amplified only 14/69 subtypes with 79.8% sensitivity (Table 2) [30].

The amplification rate for each primer pair for each subtype of HCV is shown on Table 2. The overall percentage of each pair of primers is shown at the bottom of Table 2. Based on the amplified size and sensitivity of each primer pair, two primer pairs (red) were selected: F1/R2, whose amplification rate and size were 96.45% and 691 bp, and F2/R3, 94.59% and 825 bp. Moreover, if the subtypes of the two pairs of primers failed to amplify, for example, subtype 3b and 3c on Table 2, another two pairs of primers (F2/R2 or F3/R3) were able to amplify them as compensation (Table 2). These data indicate that a combination of PCR amplification with the four pairs of PCR primers (F1/R2, F2/R3, F2/R2 and F3/R3) could amplify all of six HCV subtypes (1117/1127) with high sensitivity for amplification.

Determination of known standard HCV genotypes by sequencing NS5B region with the new primers

The cDNAs from five known standard HCV subtypes, i.e. HCV 1a, 2a, 3b, 6a and 1b, were amplified with the ten primer pairs by RT-PCR and analysed by DNA gel electrophoresis (Fig. 4a),

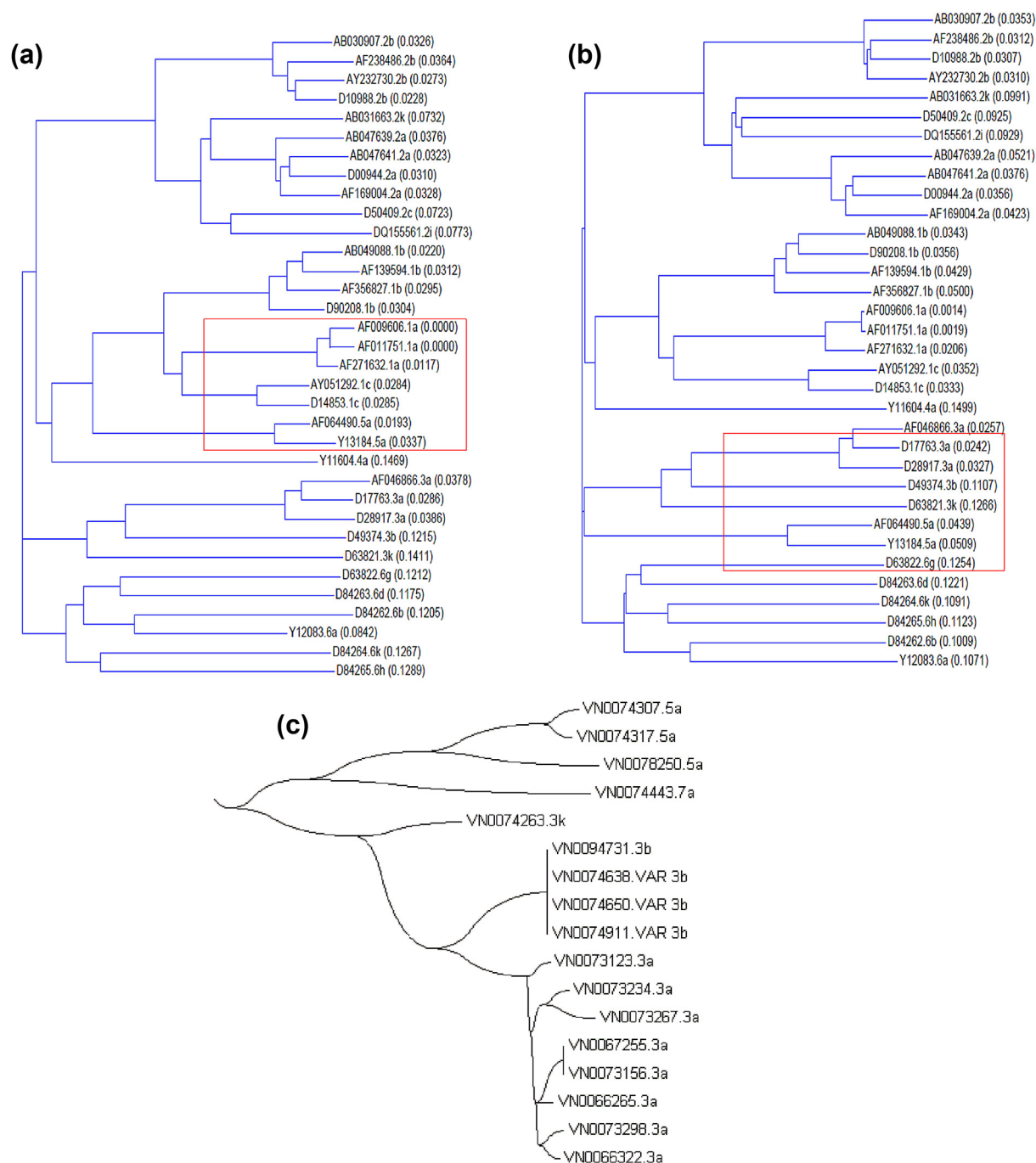


FIG. 2. Phylogenetic analysis of the genetic relatedness with different sequences of NS5B regions. (a) Phylogenetic analysis of the genetic relatedness in 32 of the nucleic acid sequences of hepatitis C virus (HCV) subtypes from the NIB database with the 8370–8708 or 8323–8638 NS5B region. (b) Phylogenetic analysis of the genetic relatedness in 32 sequences of HCV subtypes with full-length NS5B. (c) Phylogenetic analysis of the genetic relatedness in 1127 sequences of HCV subtypes with full-length NS5B.

and showed a clear band with correct size and without any non-specific bands for each primer pair; the DNA fragments were then extracted from the gel and sequenced (Fig. 4b). The resulting sequences were mapped against the HCV database (<http://www.ncbi.nlm.nih.gov/projects/genotyping>), which

showed that they are fully mapped to the correct HCV subtype (red part) (Fig. 4c). Phylogenetic analysis showed that the location of the standard NS5B sequences amplified by the primers were in the same cluster as the standard HCV stains (HCV 1a, 2a, 3b, 6a and 1b) belong to, which is indicated by a

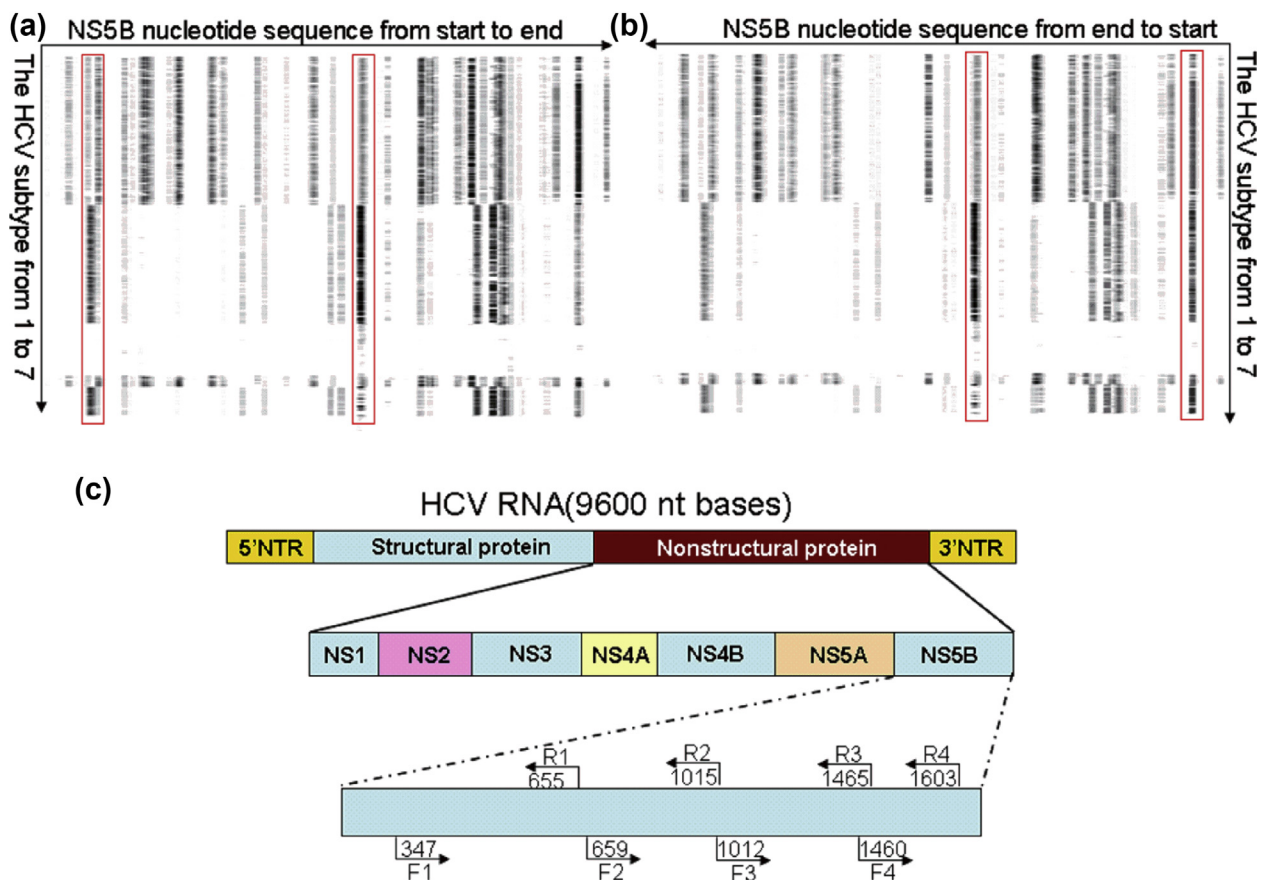


FIG. 3. Primer design and possibility for amplification of NS5B. (a, b) The potential forward (a) and reverse (b) primers for amplifying NS5B region of hepatitis C virus (HCV). Each row represents a NS5B nucleotide sequence, and each image pixel represents a base. The pixel grey-scale represents the weight of the PCR primer cluster located in NS5B region. The weight increased when the bases located in the PCR primer cluster. The deep colour region means that the region can be amplified with the PCR primers at a high amplification capacity across the whole HCV population. To clearly display pictures, the threshold value was set at 500 dpi, and the pixels were set as white colour when they were lower than the threshold. (c) Location of the potential primers in NS5B region.

red arrow in Fig. 4d. These data indicate that the predicted primers can specifically and correctly identify the standard HCV subtypes.

Detection of HCV in clinical samples by sequencing NS5B region amplified with the new primers

A total of 567 patients with HCV in the Hubei province region of China were analysed for HCV subtype by sequencing analysis using the PCR products amplified with the ten primer pairs. Fig. 5a showed the representative bands of the PCR products amplified with these primers for HCV 1b-2a. The resulting PCR products were sequenced and mapped to the HCV database. Fig. 5b is a representative result, which showed the clear sequencing and Fig. 5c showed the mapping data for HCV 1b-2a subtype. We also analysed the amplification sensitivity of each primer pair in patient samples, and found that the NS5B fragments were amplified by each of the ten primer pairs with a

high sensitivity (92.59–98.59%) (Table 3). Importantly, three new HCV subtypes (1b-2a, 1b-2k and 6d-6k) were detected in our samples that have not been previously reported in China (Table 3). We further found that the detection rate for HCV subtypes in the patients with the method, was 100%. Even the combination of only two pairs (F1/R2 and F2/R3) could amplify almost all the patient samples and if combined with another two pairs of primers (F2/R2 and F3/R3), they could amplify all the patients' samples (Table 3). These data suggested that the four pairs of primers are efficient enough for clinical subtyping of the HCV samples. We analysed the distribution ratio of the HCV subtypes in the population, and found that HCV 1b accounted for 60.3% (342/567) in the HCV patients, followed by HCV 3b and 2a, accounted for 12.0% (68/567) and 9.3% (53/567), respectively (Fig. 5d). However, the HCV 1b had a larger proportion in the previously reported Chinese population, which accounted for 41.8% (1074/2568), ranked the first; HCV

TABLE 2. The theoretical amplification ability of the potential primers for all NS5B from hepatitis C virus (HCV) database

HCV subtype	No. (n = 1127)	Primers (amplification length)										Reported [31]
		F1-R1 (328 bp)	F1-R2 (691 bp)	F1-R3 (1137 bp)	F1-R4 (1275 bp)	F2-R2 (379 bp)	F2-R3 (825 bp)	F2-R4 (963 bp)	F3-R3 (472 bp)	F3-R4 (610 bp)	F4-R4 (162 bp)	
1a	438	100.00%	97.95%	97.26%	87.67%	97.95%	97.26%	87.67%	97.26%	87.67%	87.67%	89.50%
1a-1b	6	100.00%	100.00%	100.00%	50.00%	100.00%	100.00%	50.00%	100.00%	50.00%	50.00%	100.00%
1a-2a	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
1b	355	97.75%	96.62%	96.06%	95.21%	98.03%	97.46%	96.34%	97.46%	96.34%	96.90%	83.66%
1b-2k	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
1c	3	100.00%	66.67%	100.00%	100.00%	66.67%	100.00%	100.00%	100.00%	100.00%	100.00%	66.67%
1g	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
VAR_1a	29	100.00%	96.55%	96.55%	75.86%	96.55%	96.55%	75.86%	96.55%	75.86%	75.86%	79.31%
VAR_1a_2a	4	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
VAR_1b	87	97.70%	100.00%	96.55%	96.55%	100.00%	96.55%	96.55%	96.55%	96.55%	96.55%	93.10%
VAR_1c	2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2a	19	100.00%	94.74%	100.00%	100.00%	84.21%	89.47%	89.47%	100.00%	100.00%	100.00%	78.95%
2b	25	96.00%	84.00%	88.00%	88.00%	88.00%	92.00%	92.00%	92.00%	92.00%	92.00%	80.00%
2b-1b	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2c	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2i	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2i_6p	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2k	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
VAR_2a	45	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	11.11%
3a	8	0	0	0	0	87.50%	62.50%	37.50%	62.50%	37.50%	25.00%	87.50%
3b	1	0	0	0	0	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3c	1	0	0	0	0	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
VAR_3b	3	0	0	0	0	100.00%	0	0	0	0	0	100.00%
4a	16	100.00%	100.00%	87.50%	56.25%	100.00%	87.50%	56.25%	87.50%	56.25%	56.25%	100.00%
4b	3	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
4d	3	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4f	6	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
4k	2	100.00%	100.00%	100.00%	0	100.00%	100.00%	0	100.00%	0	0	100.00%
5a	3	33.33%	33.33%	33.33%	33.33%	33.33%	66.67%	66.67%	100.00%	100.00%	100.00%	33.33%
6a	16	93.75%	100.00%	100.00%	100.00%	12.50%	12.50%	100.00%	100.00%	100.00%	100.00%	0
6b	1	100.00%	0	100.00%	100.00%	0	0	0	0	0	100.00%	0
6c	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
6d	1	100.00%	100.00%	100.00%	100.00%	0	0	0	100.00%	100.00%	100.00%	0
6e	4	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	50.00%
6f	3	33.33%	100.00%	100.00%	0	66.67%	66.67%	0	100.00%	0	0	0
6g	3	100.00%	100.00%	100.00%	100.00%	66.67%	66.67%	66.67%	100.00%	100.00%	100.00%	33.33%
6h	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6i	3	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
6j	2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6k	5	20.00%	20.00%	20.00%	20.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
6l	2	0	0	0	0	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	50.00%
6m	4	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
6n	5	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	20.00%
6o	2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0
6p	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6q	1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6t	4	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
7a	1	0	0	0	0	0	0	0	0	0	100.00%	0
Overall	1127	96.45%	95.12%	94.59%	88.91%	95.56%	94.59%	88.73%	96.36%	90.42%	90.68%	79.86%

3b (12.5%) and 2a (10.2%) ranked third and fourth, respectively (Fig. 5e). HCV 6a subtype accounted for 18.0% in the reported Chinese population (462/2568), ranked the second, but it only ranked the fourth (2.1%) in our samples, far lower than that in the reported. The detective rate for new HCV subtypes in the samples was HCV 1b-2a 2.8%, 1b-2k 2.1%, and 6d-6k 2.3% (Fig. 5D). These data indicates this method is more sensitive for subtyping of clinical HCV samples.

Discussion and conclusions

There are many methods for classifying HCV by analysis of certain regions of the virus based on the length and diversity of the HCV genome; however, the protocols still lack

standardization [31]. Here, we developed a more accurate and more reliable method for genotyping HCV by sequencing NS5B regions amplified with the primer pair from the eight identified primers. This method has higher sensitivity for amplification and 100% sensitivity for subtyping of clinical HCV samples.

Sequence analysis of NS5B for genotyping HCV provides precise genotype and subtype identification and an accurate epidemiological representation of circulating viral strains. However, current HCV NS5B-based genotyping uses only partial short nucleic acid sequences of NS5B. Kozletski et al. [32] used a 329-bp NS5B gene fragment (nucleotides 8280–8610 in HCV sequence) to classify HCV and showed an 89% clinical sensitivity for amplification. Cantaloube et al. [30] used 339-bp nucleotide segments (nucleotides 8370–8708) in the NS5B region to classify HCV and achieved a 98.32% of sensitivity for

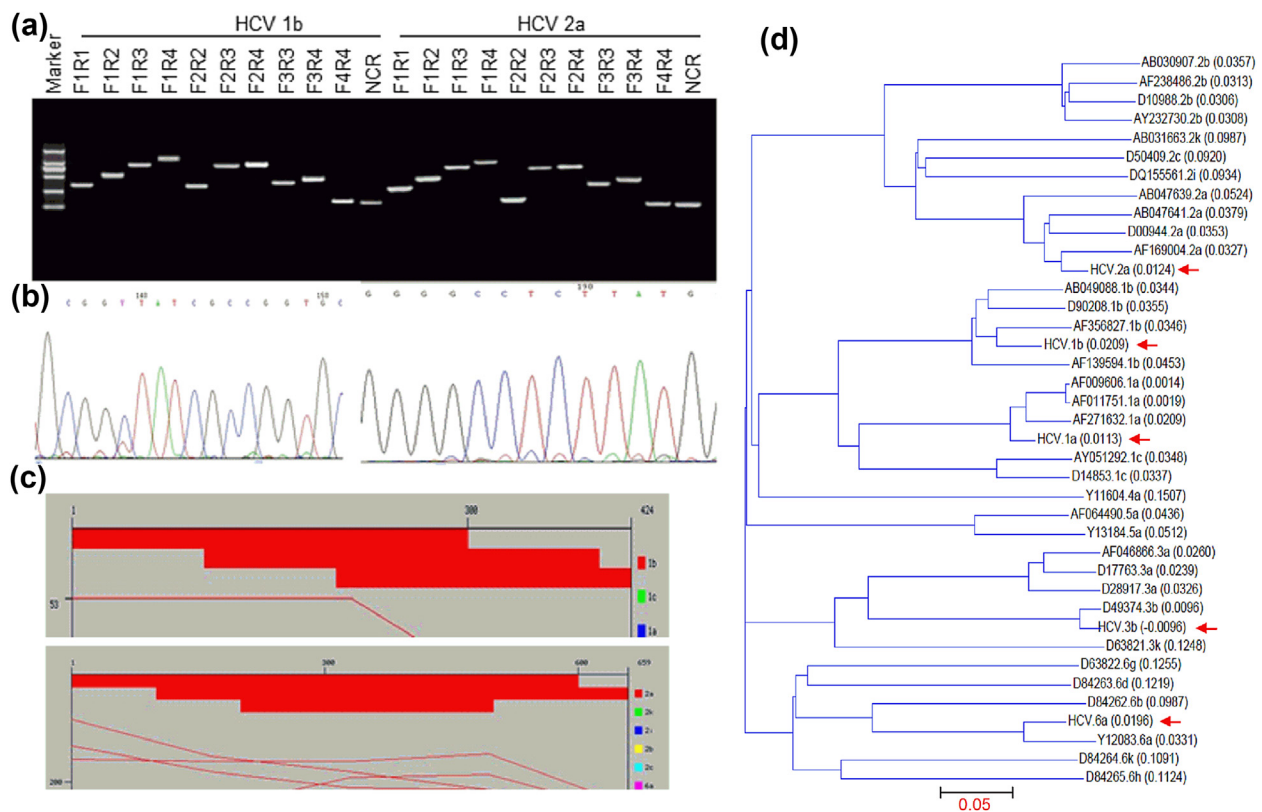


FIG. 4. Amplification of the standard hepatitis C virus (HCV) subtypes with the predicted primers. (a) Gel electrophoresis after PCR amplification of HCV 1b and 2a. (b) Sequencing results of the PCR products of HCV 1b and 2a. (c) Mapping the HCV 1b and 2a sequences with the HCV database. The red part indicates that they are fully aligned. (d) Phylogenetic analysis of the HCV 1b and 2a sequences and their location in phylogenetic tree of standard HCV stains HCV 1a, 2a, 3b, 6a and 1b (see the red arrow) using CLUSTALW2 software.

amplification; however, they only detected HCV type 4 and 1. Murphy *et al.* [33] used primers located in the NS5B gene (nucleotide 8323–8638) to classify HCV and reached a 97.3% sensitivity for amplification with many random primers. However, our study showed that it could not distinguish between subtype 5 and 1 or between 4 and 5 if only sequencing the reported nucleotide segment (nucleotide 8370–8708 or 8323–8638) in the NS5B region (Fig. 2a). Our method here is designed to sequence NS5B region amplified with only eight primers, which has theoretically 99.20% (1117/1127) sensitivity for amplification, and is able to detect six of seven genotypes of HCV, and even reach 100% sensitivity for HCV subtyping in our observed clinical samples.

The major characteristic of our method is the use of a computer tool to search for the best primers for sequencing the NS5B region using the entire NS5B sequence from all 1127 HCV genomic DNAs and the eight best primers. The computer tool is also used to examine the possibility of amplification using the ten primer pairs and it found that four primer pairs (F1/R2, F2/R3, F2/R2 and F3/R3) could amplify all the available HCV genomic

sequences in six of seven HCV genotypes. Our result also showed that this method could detect the five known standard HCV subtypes with 100% accuracy. More importantly, this method could detect the subtypes of all 567 clinical HCV samples with almost 100% sensitivity and even only using the four pairs of primers (F1/R2, F2/R3, F2/R2 and F3/R3). Of course, the best genotyping method is sequencing the whole HCV genome; however, this is not applicable for clinical samples. Our results indicated that sequencing the amplified NS5B region with our identified primers is a reliable, sensitive and economical method for clinical genotyping and subtyping of the HCV.

Correctly genotyping and subtyping HCV is critical for clinical diagnosis and treatment of HCV, and also HCV classification and epidemiology studies [34]. Our subtyping data showed that the subtypes of HCV 1b, 2a, 3b and 6a, accounted for 60.35%, 9.3%, 12.0% and 2.1%, respectively in our observed patient samples, whereas the frequency of HCV subtyping in the previously reported Chinese population (2010), HCV 1b, 2a, 3b and 6a, accounted for 41.8%, 10.2%, 12.5% and 18.0%, respectively. The percentage of HCV 1b in our samples is much

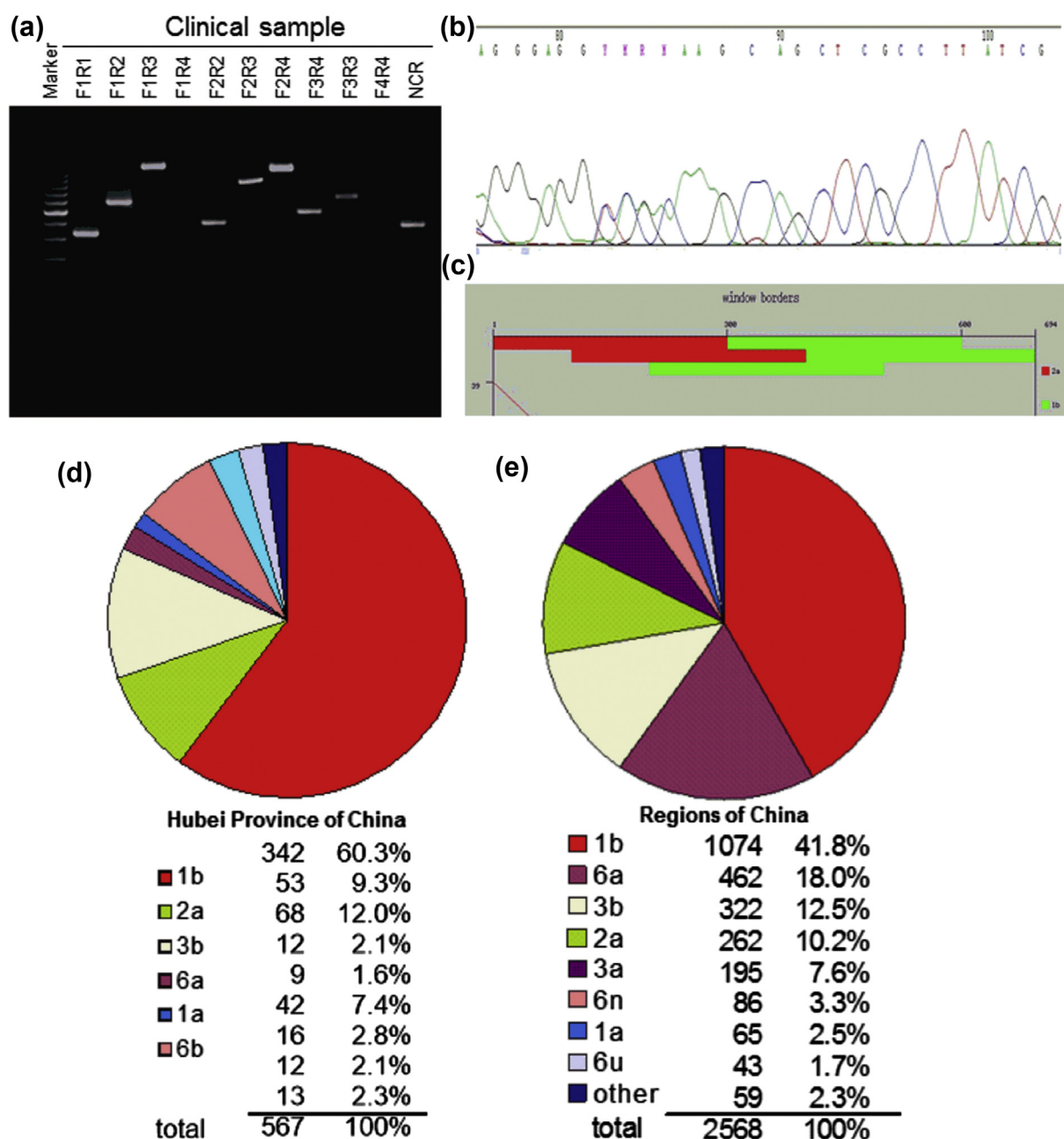


FIG. 5. Distribution of the hepatitis C virus (HCV) genotypes identified in the HCV patients. (a) Gel electrophoresis after PCR amplification of HCV 1b-2a. (b) Sequencing results of the PCR products of HCV 1b-2a. (c) Mapping the HCV 1b-2a sequence with the HCV database. (d) Percentage of HCV genotypes identified in the population of Hubei province of China. (e) Percentage of HCV genotypes identified in all Chinese population from the HCV database.

higher than previously reported. We also compared the prevalence of our HCV genotypes with that of worldwide reports. Recently, Gower *et al.* [35] and Messina *et al.* [36] reported that HCV genotype 1 was the most common (46% or 46.2%) in HCV patients of 87 or 117 countries, followed by genotype 3 (22% and 30.1%). Both genotype 2 and 4 occupied 13% in Gower *et al.*'s report and Messina *et al.* also observed

genotypes 2, 4 and 6 responsible for a total 22.8% of all cases [35,36]. Our data showed that the genotype 1 (1b + 1a) is dominant genotype of HCV in China, as worldwide, occupied 61.9% (60.3 + 1.6%) of total patients in the cohort, followed by genotype 3 (3b, 12%), genotype 6 (6a + 6b, 9.5%) and genotype 2 (2a, 9.3%). These data also showed a higher rate for genotype 1 (mainly 1b) than that found in worldwide reports.

TABLE 3. The amplification ability of the potential primers for 567 patients with hepatitis C virus (HCV)

HCV subtype	No. (n = 567)	Primers (amplification length)									
		F1-R1 (328 bp)	F1-R2 (691 bp)	F1-R3 (1137 bp)	F1-R4 (1275 bp)	F2-R2 (379 bp)	F2-R3 (825 bp)	F2-R4 (963 bp)	F3-R3 (472 bp)	F3-R4 (610 bp)	F4-R4 (162 bp)
1a	9	100.00%	100.00%	88.89%	88.89%	100%	100.00%	77.78%	77.78%	66.7%	66.67%
1b	342	97.66%	97.66%	97.37%	97.95%	98.54%	99.71%	92.40%	96.78%	94.74%	90.94%
1b-2k	12	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2a	53	100.00%	92.45%	100.00%	100.00%	94.34%	98.11%	90.57%	100.00%	100.00%	100.00%
3b	68	97.06%	86.76%	91.18%	77.94%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6b	42	100.00%	100.00%	100.00%	100.00%	88.10%	85.71%	78.57%	90.48%	92.65%	100.00%
1b-2a	16	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6d-6k	13	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6a	12	91.67%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Overall	Sensitivity	98.06%	96.30%	97.19%	95.94%	97.71%	98.59%	92.59%	97.00%	95.41%	94.00%

Hepatitis C virus genotype 1b was reported to be associated with progression to chronicity, cirrhosis, and severe and decompensated liver disease [37,38]; also rapid recurrence and severe hepatitis in transplant patients when compared with the other HCV genotypes [39]. A linkage of HCV 1b with the frequency and development of HCC was also reported [40–43]. These studies suggested that genotype 1b is a potential marker for more severe HCV-associated liver disease. We observed an increase of HCV 1b, which not only suggested the sensitivity increase of our method, but also suggested an increase of HCV-infection-associated severe liver disease in China.

Moreover, Petruzzello *et al.* [44] reported that the rate of genotype 1 in Italy occupied 63.6% of total HCV cases, followed by genotype 2 (29.4%), genotype 3 (6.2%) and genotype 4 (0.8%). They also found that subtype 1b was more frequent in females than in males. Conversely, genotype 3 was more frequent in males [44]. The reported rate of genotypes 1 and 3 is similar to that of our data. However, the percentage of genotype 2 is much higher than our data; also we did not observe the sex difference of genotypes 1b and 3 in our cohort (data not shown).

Hepatitis C virus genotype 6 (HCV 6) is restricted to South China, South East Asia, and is also occasionally found in migrant patients from endemic countries. HCV 6 has considerable genetic diversity with 23 subtypes (a to w) [45]. Our data showed that the frequency of HCV 6a is dramatically reduced in our observed samples compared with previous reports (Table 2). The reason for this difference may be because our sequencing method is more specific and more reliable for subtyping, which prevents other subtypes also mistakenly classifying as HCV 6a.

We did not detect genotypes 4 and 5, which have the largest proportions in lower-income areas. Instead, we identified three new HCV subtypes 1b-2a, 1b-2k and 6d-6k. HCV was thought to evolve in a clonal manner, with diversity generated through the accumulation of mutations. However, several events of inter- and intragenotypic homologous recombination (1a-1b, 1a-1c, 2b-1b, 2k-1b, 2-5, 1b-1a, 2i-6h) have been reported to date from different regions of the world [46–48]. We believe

that the identified new subtypes result from putative inter- and intragenotypic recombination. Also, the recombination events were detected in 25 of the 234 independent samples analysed (10.7%) [49]. We observed that the detection rates for the new subtypes were HCV 1b-2a 2.8%, 1b-2k 2.1% and 6d-6k 2.3% in the NS5A region, which is lower than the reports. Only a few studies have reported the recombination in natural populations of HCV [49]. Our study is the first report of the putative recombination of HCV in Chinese patients. The recombination is considered as a potentially relevant mechanism generating genetic variation in HCV with important implications for the treatment of this infection [49]. Unfortunately, most of the HCV-infected patients in this study were outpatients. Detailed clinical data for the new HCV subtypes were lacking. Patients in this study included those with acute hepatitis C, chronic hepatitis C, hepatitis C cirrhosis and HCV acute exacerbation. We are unable to make a definite correlation between the new HCV subtypes and a specific appearance of clinical conditions; however, we believe that the new HCV subtypes may appear mainly in the patients with severe hepatitis or acute exacerbation and reactivation of chronic HCV infection or HCC development; and the clinical significance of the new HCV subtypes will be further clarified.

In conclusion, the new NS5B regions amplified with the newly designed primers can be used to genotype HCV with high sensitivity for amplification, higher accuracy and 100% sensitivity for subtyping clinical HCV samples. With this method, we observed changes in the ratio of HCV subtypes and identified new HCV subtypes, which is helpful for clinical diagnosis, treatment, classification of HCV infection and the study of HCV epidemiology and vaccine development.

Transparency declaration

All authors report that there are no conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of

Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

Acknowledgements

We thank Melanie Hartman and Sadie Steffens for English grammar correction. This work supported by the Hubei Provincial Natural Science Foundation of China (2010CDB06903), National Natural Science Foundation of China (81000771), National Key Basic Research Programme of China – 973 Programme (2012CB526706) and the National Natural Science Foundation of China (NSFC81271694).

References

- [1] Thomas DL. Global control of hepatitis c: where challenge meets opportunity. *Nat Med* 2013;19:850–8.
- [2] Mohd Hanafiah K, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 2013;57:1333–42.
- [3] Lavanchy D. Evolving epidemiology of hepatitis c virus. *Clin Microbiol Infect* 2011;17:107–15.
- [4] Venook AP, Papandreou C, Furuse J, de Guevara LL. The incidence and epidemiology of hepatocellular carcinoma: a global and regional perspective. *Oncologist* 2010;15(Suppl. 4):5–13.
- [5] Negro F, Alberti A. The global health burden of hepatitis C virus infection. *Liver Int* 2011;31(Suppl. 2):1–3.
- [6] Simmonds P. Viral heterogeneity of the hepatitis C virus. *J Hepatol* 1999;31(Suppl. 1):54–60.
- [7] Irshad M, Ansari MA, Singh A, Nag P, Raghvendra L, Singh S, et al. HCV-genotypes: a review on their origin, global status, assay system, pathogenicity and response to treatment. *Hepato-gastroenterology* 2010;57:1529–38.
- [8] Smith DB, Simmonds P. Hepatitis C virus: types, subtypes, and beyond. *Methods Mol Med* 1999;19:133–46.
- [9] Kuiken C, Yusim K, Boykin L, Richardson R. The Los Alamos hepatitis C sequence database. *Bioinformatics* 2005;21:379–84.
- [10] Kuiken C, Hraber P, Thurmond J, Yusim K. The hepatitis C sequence database in Los Alamos. *Nucleic Acids Res* 2008;36:D512–6.
- [11] Dusheiko G, Simmonds P. Sequence variability of hepatitis C virus and its clinical relevance. *J Viral Hepat* 1994;1:3–15.
- [12] Mondelli MU, Silini E. Clinical significance of hepatitis C virus genotypes. *J Hepatol* 1999;31(Suppl. 1):65–70.
- [13] Croagh CM, Lubel J. Advances in the management of hepatitis C. *Intern Med J* 2013;43:1265–71.
- [14] Nelson PK, Mathers BM, Cowie B, Hagan H, Des Jarlais D, Horyniak D, et al. Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. *Lancet* 2011;378:571–83.
- [15] Grebely J, Dore GJ. What is killing people with hepatitis C virus infection? *Semin Liver Dis* 2011;31:331–9.
- [16] Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. *Nat Rev Gastroenterol Hepatol* 2013;10:553–62.
- [17] Nguyen LH, Nguyen MH. Systematic review: Asian patients with chronic hepatitis C infection. *Aliment Pharmacol Ther* 2013;37:921–36.
- [18] Gryadunov D, Nicot F, Dubois M, Mikhailovich V, Zasedatelev A, Izopet J. Hepatitis C virus genotyping using an oligonucleotide microarray based on the NSSB sequence. *J Clin Microbiol* 2010;48:3910–7.
- [19] Larrat S, Poveda JD, Coudret C, Fusillier K, Magnat N, Signori-Schmuck A, et al. Sequencing assays for failed genotyping with the versant hepatitis C virus genotype assay (lipa), version 2.0. *J Clin Microbiol* 2013;51:2815–21.
- [20] Sam SS, Steinmetz HB, Tsongalis GJ, Tafe LJ, Lefferts JA. Validation of a solid-phase electrochemical array for genotyping hepatitis C virus. *Exp Mol Pathol* 2013;95:18–22.
- [21] Nakao T, Enomoto N, Takada N, Takada A, Date T. Typing of hepatitis C virus genomes by restriction fragment length polymorphism. *J Gen Virol* 1991;72(Pt 9):2105–12.
- [22] Weck K. Molecular methods of hepatitis C genotyping. *Exp Rev Mol Diagn* 2005;5:507–20.
- [23] Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* 2014;59:318–27.
- [24] Laperche S, Lunel F, Izopet J, Alain S, Dény P, Duverlie G, et al. Comparison of hepatitis C virus NSSB and 5' noncoding gene sequencing methods in a multicenter study. *J Clin Microbiol* 2005;43:733–9.
- [25] Nakano T, Lau GM, Sugiyama M, Mizokami M. An updated analysis of hepatitis C virus genotypes and subtypes based on the complete coding region. *Liver Int* 2012;32:339–45.
- [26] Golemba MD, Culasso AC, Villamil FG, Bare P, Gadano A, Ridruejo E, et al. Hepatitis C virus diversification in Argentina: comparative analysis between the large city of Buenos Aires and the small rural town of O'Brien. *PloS One* 2013;8:e84007.
- [27] Nakatani SM, Santos CA, Riediger IN, Krieger MA, Duarte CA, do Carmo Debur M, et al. Comparative performance evaluation of hepatitis C virus genotyping based on the 5' untranslated region versus partial sequencing of the NSSB region of Brazilian patients with chronic hepatitis C. *Virol J* 2011;8:459.
- [28] Elsadek Fakhr A, Pourkarim MR, Maes P, Atta AH, Marei A, Azab M, et al. Hepatitis C virus NSSB sequence-based genotyping analysis of patients from the Sharkia governorate. *Egypt Hepat Monthly* 2013;13:e12706.
- [29] Simmonds P, Mellor J, Sakuldamrongpanich T, Nuchaprayoon C, Tanprasert S, Holmes EC, et al. Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. *J Gen Virol* 1996;77(Pt 12):3013–24.
- [30] Cantaloube JF, Laperche S, Gallian P, Bouchardeau F, de Lamballerie X, de Micco P. Analysis of the 5' noncoding region versus the NSSB region in genotyping hepatitis C virus isolates from blood donors in France. *J Clin Microbiol* 2006;44:2051–6.
- [31] Opreșan G, Szmal C, Dinu S, Opreșoreanu AM, Thiers V, Panait M, et al. Comparative methods for genotyping hepatitis C virus isolates from Romania. *Roum Arch Microbiol Immunol* 2009;68:151–7.
- [32] Koletzki D, Dumont S, Vermeiren H, Fevery B, De Smet P, Stuyver LJ. Development and evaluation of an automated hepatitis C virus NSSB sequence-based subtyping assay. *Clin Chem Lab Med* 2010;48:1095–102.
- [33] Murphy DG, Willems B, Deschenes M, Hilzenrat N, Mousseau R, Sabbah S. Use of sequence analysis of the NSSB region for routine genotyping of hepatitis C virus with reference to c/e1 and 5' untranslated region sequences. *J Clin Microbiol* 2007;45:1102–12.
- [34] Jacka B, Lamoury F, Simmonds P, Dore GJ, Grebely J, Applegate T. Sequencing of the hepatitis C virus: a systematic review. *PloS One* 2013;8:e67073.
- [35] Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J Hepatol* 2014;61:S45–57.

- [36] Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, et al. Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology* 2015;61:77–87.
- [37] Zein NN. Clinical significance of hepatitis C virus genotypes. *Clin Microbiol Rev* 2000;13:223–35.
- [38] Tanaka H, Tsukuma H, Yamano H, Okubo Y, Inoue A, Kasahara A, et al. Hepatitis C virus 1b(ii) infection and development of chronic hepatitis, liver cirrhosis and hepatocellular carcinoma: a case–control study in Japan. *J Epidemiol* 1998;8:244–9.
- [39] Sugo H, Balderson GA, Crawford DH, Fawcett J, Lynch SV, Strong RW, et al. The influence of viral genotypes and rejection episodes on the recurrence of hepatitis C after liver transplantation. *Surg Today* 2003;33:421–5.
- [40] Murase J, Kubo S, Nishiguchi S, Hirohashi K, Shuto T, Ikebe T, et al. Correlation of clinicopathologic features of resected hepatocellular carcinoma with hepatitis C virus genotype. *Jpn J Cancer Res* 1999;90:1293–300.
- [41] Sobesky R, Feray C, Rimlinger F, Derian N, Dos Santos A, Roque-Afonso AM, et al. Distinct hepatitis C virus core and f protein quasispecies in tumoral and nontumoral hepatocytes isolated via microdissection. *Hepatology* 2007;46:1704–12.
- [42] Tamura M, Yamada A, Takao Y, Matsueda S, Komatsu N, Yutani S, et al. Identification of hepatitis C virus 1b-derived peptides recognized by both cellular and humoral immune systems in HCV1b⁺ HLA-A24⁺ patients. *Hepatol Res* 2005;32:227–34.
- [43] Jiang J, Wu X, Tang H, Luo G. Apolipoprotein E mediates attachment of clinical hepatitis C virus to hepatocytes by binding to cell surface heparan sulfate proteoglycan receptors. *PLoS One* 2013;8:e67982.
- [44] Petruzzello A, Coppola N, Loquercio G, Marigliano S, Giordano M, Azzaro R, et al. Distribution pattern of hepatitis C virus genotypes and correlation with viral load and risk factors in chronic positive patients. *Intervirology* 2014;57:311–8.
- [45] Thong VD, Akkarathamrongsin S, Poovorawan K, Tangkijvanich P, Poovorawan Y. Hepatitis C virus genotype 6: virology, epidemiology, genetic variation and clinical implication. *World J Gastroenterol* 2014;20:2927–40.
- [46] Kageyama S, Agdamag DM, Alesna ET, Leñaño PS, Heredia AM, Abellanos-Tac-An IP, et al. A natural inter-genotypic (2b/1b) recombinant of hepatitis C virus in the Philippines. *J Med Virol* 2006;78:1423–8.
- [47] Moreno MP, Casane D, Lopez L, Cristina J. Evidence of recombination in quasispecies populations of a hepatitis C virus patient undergoing anti-viral therapy. *Virology* 2006;3:87.
- [48] Legrand-Abravanel F, Claudinon J, Nicot F, Dubois M, Chapuy-Regaud S, Sandres-Saune K, et al. New natural intergenotypic (2/5) recombinant of hepatitis C virus. *J Virol* 2007;81:4357–62.
- [49] Sentandreu V, Jiménez-Hernández N, Torres-Puente M, Bracho MA, Valero A, Gosalbes MJ, et al. Evidence of recombination in inpatient populations of hepatitis C virus. *PLoS One* 2008;3:e3239.